PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The relationship between wealth and emotional well-being before, during, versus after a nationwide disease outbreak: A large-scale investigation of disparities in psychological vulnerability across COVID-19 pandemic phases in China
AUTHORS	Yang, Haiyang; Ma, Jingjing

VERSION 1 – REVIEW

REVIEWER	david Blanchflower Department of Economics Dartmouth College Hanover NH 03755 USA
REVIEW RETURNED	30-Sep-2020

GENERAL COMMENTS	Review of Disparity in psychological vulnerability: The relationship between wealth and emotional well-being before, during, versus after the nationwide COVID-19 outbreak in China by Yang and Ma This paper explores an interesting question as to who was vulnerable, in emotional terms, to the COVID19 shock in China. The paper finds, unsurprisingly that income and wealth provide buffers against the shock. I would have predicted that. There is a large literature that already establishes that education, income, and wealth raise happiness, both at the individual and macro levels – see work by Easterlin and Graham for example. A major omission in the paper would also seem to be that previous health conditions would also likely make individuals more vulnerable – examples would be presence of asthma, COPD, diabetes etc. Rich people are healthier and live longer. Is it health or wealth that is driving the results? This has also been documented recently in the US and elsewhere. The originality here is the data. I have some problems with what has been done. 1. The dependent variable is a conglomerate of answers from questions on whether an individual smiled or laughed yesterday and whether they experienced a lot of enjoyment/ happiness/anger/sadness/stress/ worry yesterday. Apparently, the authors "subtracted the average an of the negative emotions experienced from the average of the positive emotions experienced". Why that is done is not explained. So, a happy person says yes to smile or laughed/happy enjoyed so scores a 3

	and zero to the four negative affect variables anger/sadness/stress/worry. My guess is the dependent variable can take the value from +3 to -4? Of interest is how that distribution changed before during and after. Which of the seven was driving the action? 2. We are not told anything about the distribution of this variable, its mean or how it changed over time. What if the only thing that changes over time is anger or two rise and two fall so overall there is no change? How many people are unhappy? 3. The question is whether the results are the same if each of these variables are examined individually? At least this way non-response bias will be less of an issue given anyone who fails to answer any one of the seven variables will be omitted from the author's analysis and this bias is unlikely to be random. Why not just sum up the positive affect variables and examine that and do the same for the negative affect variables? As written, I have no idea what is going on. 4. No sample sizes are given in the tables. Should I assume the sample size is less than the three samples of 11,131; 3,000 and 13,269 described in the methods section? 5. It seems very surprising that none of the individual characteristic variables in Tables 1 & 2 are significant. Examples are age, gender, marital status. An obvious omission of course is education and labor force status. I think this is likely a reflection of the strangeness of the dependent variable. The expectation is they would enter significantly if each of the variables were estimated individually as I suggested above. 6. The figures don't help much especially for the Emotional Well-being variable that apparently is scored from zero to 1 in Figure 1? But if as he authors explain the four unhappy scores are deducted from the three happy ones then an unhappy person should have a score of -4 and a happy person a score of +3. I understand the other scores will have means between 0 and 1.

REVIEWER	Ronald Fischer Instituto D'Or of Research & Teaching; Brazil Victoria University of Wellington
REVIEW RETURNED	26-Nov-2020

GENERAL COMMENTS	This is an important analysis and dataset. Issues to address: Describe your data and model in more detail. The same individuals are measured at 3 time points? If yes, carefully address the lower response rate at T2 - are non-responders at T2 (and possibly T3) different from responders at T1? Include a test whether the results differ between Hubei vs elsewhere. I strongly recommend testing the structure of the well-being scales before, during and after: e.g., a test of the temporal stability/invariance over time. If the measures are not temporally invariant, this needs to be reported and the results need to be interpreted with caution. Many within-person methods assume temporal invariance. Examining the plots, it appears that the pattern may differ by item/valence vs arousal components.

	Do not use difference scores and rather use appropriate composite scores. Previous research has shown that positive and negative mood effects are not identical. Furthermore, difference scores are statistically problematic (see the extended work by Edwards on the problems with difference scores). In the current context, a separate analysis for valence (possibly arousal) would be most informative. I wonder whether it would be useful to either use latent growth models or random effects regression with random effects for individuals. The analysis at this moment is not clearly specified to allow interpretation of the results. Report effective N for the analyses (if necessary by wave, after matching). I would recommend plotting the results over time, separate lines by ownwership/income. It may also be useful to use more appropriate within-person visualizations (for some promising examples, see https://github.com/jorvlan/open-visualizations).

VERSION 1 – AUTHOR RESPONSE

Response to Reviewer 1 (Professor David Blanchflower)

Thank you very much for the opportunity to revise our manuscript. We greatly appreciate the time and effort you put into reviewing our paper. We are grateful for your helpful suggestions. We have revised the manuscript following your guidance.

You asked us to better explicate the emotional well-being measure used, and present the distributions of the emotions. In the revised paper, we more clearly noted in the paper that the overall emotional well-being index was created by subtracting the average of the negative emotions (which had a value range of 0-1) from the average of the positive emotions (which had a value range of 0-1). Thus, the overall index had a value range of 0-1, and accounted for experiences of both positive and negative emotions. (Prior research [e.g., the work by Diener and colleagues] utilized similar overall indexes.) More importantly, to illustrate the distributions of the emotions across the three time periods, we created detailed tables laying out the percentage of individuals experiencing each of the emotions during each phase (see tables 1a and 1b). We believe that these new additions, along with figures 1 and 2, clearly show the changes in each emotion over the pandemic phases.

You also suggested that we construct an index for “the positive affect variables and examine that and do the same for the negative affect variables.” Following your guidance, we created an index of positive emotions as well as one for negative emotions. Analyses using these two indexes yielded patterns of results largely consistent with the results using the overall emotional well-being index. We rewrote the results section (section 3) to present the findings of all three indexes.

You asked about the final sample sizes in our analyses. In the revised paper, we more clearly specified the sample size for every analysis we conducted (please see “N” in table 1a, 1b, 2, 3, & 4).

You noted about the potential effects of demographic variables, and suggested education and labor force status as additional control variables. In this research, the data available to us did not allow us to investigate the potential effect of education and labor force status across the three periods. However, as you predicted, in regression analyses with the positive emotion index or negative emotion index as the dependent variable, the coefficient estimates of demographic variables such as age, gender, and marital status were significant (see tables 3 and 4). We discussed these additional findings in the results section.

You also noted that we should more clearly explicate the range of the overall emotional well-being index. To address this, we added an explanation of the value range in the methods section.

Finally, we also followed up on your comment regarding the potential role of physical health in driving the results observed. First, in our analyses, we controlled for the effect of age, which often correlates with physical health status. The analyses thus suggest that our findings might not be simply driven by differences in physical health. Second, we obtained a dataset on individuals' perception of their physical health (https://opendata.pku.edu.cn/dataverse/CFPS) and merged it with our own data. Specifically, given the available fields of the two datasets, we first calculated an index of health level per location (i.e., province) using the CFPS dataset. We then merged that index into our dataset, matching on location. Hence, for each record, we were able to control for the physical health level of the location. We re-ran our analyses controlling for the health level. The analyses yielded the same patterns of significant results, suggesting that our findings are driven by wealth. Nonetheless, we completely agree with you that further research is needed to conclusively disentangle the effect of physical health from that of wealth. We have included a discussion about this in section 4 of the manuscript.

Thank you again for your insightful comments, which helped improve our research substantially.

Response to Reviewer 2 (Professor Ronald Fischer)

Thank you very much for the opportunity to revise our manuscript. We greatly appreciate the time and effort you put into reviewing our paper. We are grateful for your helpful comments and suggestions. We have revised the manuscript following your guidance.

You suggested that we more clearly describe whether our datasets and models are within-participant in nature. We clarified these in the methods and results sections of the revised manuscript. Specifically, we noted that the data for each pandemic period were collected using the same sampling method (as opposed to measuring the same individuals repeatedly). In other words, the data and hence the models we used are cross-sectional in nature. Furthermore, we also more clearly specified the sample size for every analysis we conducted (please see "N" in table 1a, 1b, 2, 3, & 4).

You noted about whether the dummy variable Hubei exhibited a significant effect. In all analyses (tables 2-4) where Hubei was included as a control variable, the coefficient estimate for that dummy was not significant. We also conducted additional analyses to explore whether the dummy variable might interact with wealth (income and real estate ownership) – the focal construct of our investigation. Those interactions were also not significant.

You suggested that, if our data were within-participant in nature, we conduct within-participant analyses and plot the results accordingly. We very much appreciate these suggestions. Given that the nature of our datasets (i.e., three cross-sectional datasets collected using the same sampling method), we were not able to utilized within-participant analyses (e.g., models with a random coefficient for each individual) or plotting approaches. However, we completely agree with you that it would be important to examine the phenomenon using such approaches and noted about this in discussion section of the manuscript.

You suggested that we conduct our analyses using an index of positive emotions and one for negative emotions. Following your guidance, we constructed these indexes and included respective analyses. We agree with you that there are limitations to an overall emotional well-being index. As some prior research utilized similar indexes (e.g., the work by Diener and colleagues), we ended up

retaining the overall index. We explained in the revised manuscript that this index is a single measure accounting for experiences of both positive and negative emotions. More importantly, analyses using the positive and negative emotion indexes yielded patterns of results largely consistent with the results using the overall emotional well-being index. We rewrote the results section (section 3) to present the findings of all three indexes.

Thank you again for your insightful comments, which helped improve our research substantially.